Received: 24 September 2018 | Revised: 18 December 2018 | Accepted: 5 January 2019 | Published on: 19 February 2019

DOI: 10.1111/aab.12491



Annals of Applied Biology WILEY

RESEARCH ARTICLE

# Genetic diversity and population structure of *Pepino mosaic virus* in tomato crops of Spain and Morocco

Cristina Gómez-Aix[1*] | Cristina Alcaide[2*] | Jesús Agüero[1] | Mohamed Faize[3] | Miguel Juárez[4] | Carlos J. Díaz-Marrero[1] | Moises Botella-Guillén[5] | Ana I. Espino[5] | Miguel A. Aranda[2] | Pedro Gómez[2]

[1]Abiopep S.L. Parque Científico de Murcia, Ctra. de Madrid Km 388, Murcia, Spain

[2]Centro de Edafología y Biología Aplicada del Segura (CEBAS)-CSIC, Departamento de Biología del Estrés y Patología Vegetal, Murcia, Spain

[3]Laboratory of Plant Biotechnology, Ecology and Ecosystem Valorisation, Faculty of Sciences, University Chouaib Doukkali, El Jadida, Morocco

[4]Escuela Politécnica Superior de Orihuela, Universidad Miguel Hernández de Elche, Alicante, Spain

[5]Plant Health Laboratory of Tenerife, Gobierno de Canarias, Islas Canarias, Spain

**Correspondence**
Pedro Gómez, Centro de Edafología y Biología Aplicada del Segura (CEBAS)-CSIC, Departamento de Biología del Estrés y Patología Vegetal, 30100 Murcia, Spain.
Email: pglopez@cebas.csic.es

**Funding information**
ERANET-ARIMNet2, Grant/Award Number: PCIN-2017-055 ; Ministry of Economy, Industry and Competitiveness, Grant/Award Number: AGL2014-59556-R; FPU16/02569; PTQ-13-05882; PTQ-15-07646; Ministère de l'Enseignement Supérieur, de la Recherche Scientifique de la Formation des Cadres

*Pepino mosaic virus* (PepMV, genus *Potexvirus*) is an emergent and highly infectious pathogen responsible for economically important diseases in tomato crops. An extensive survey of tomato plants showing PepMV-like symptoms was carried out in 2017 to study the PepMV genetic diversity and populations structure in different tomato-producing areas of Spain and Morocco. Molecular dot-blot hybridization analysis showed that virus populations from Spain and Morocco were mainly composed of isolates belonging to the Chilean 2 (CH2) strain, although isolates of the European (EU) strain were detected in significant proportions in Spanish populations, mainly in mixed infections. A few isolates of the American (US1) strain were also detected in Tenerife (Canary Islands, Spain) crops. Eighty-five isolates were randomly selected and sequenced in the genomic region that encodes the triple gene block and capsid protein genes. Our phylogenetic and population genetics analyses confirmed the presence of the CH2, EU and US1 PepMV strains. Despite the high genetic similarity observed within populations, variants were maintained at low frequency under purifying selection, and differentiation among more geographically distant locations was identified, with potential gene flow contributing to the shaping of the PepMV populations structure.

**KEYWORDS**

geographical distribution, mixed genotype infections, PepMV, population structure, tomato crop, viral genetic diversity

## 1 | INTRODUCTION

Emerging infectious diseases in crops are a serious threat to agriculture and global food sustainability (Jones, 2009; Oerke, 2005; Vurro, Bonciani, & Vannacci, 2010; Woolhouse, Haydon, & Antia, 2005). Plant viruses cause almost half of the reported emerging diseases (Anderson et al., 2004), and among them, those induced by RNA plant viruses are of particular concern due to their overrepresentation and the lack of effective countermeasures. The emergence and expansion of plant viruses have been attributed to agricultural intensification, global trade of seeds and products and the introduction and expansion of insect vectors (e.g., *Bemisia tabaci*) (Anderson et al., 2004; Hanssen, Lapidot, & Thomma, 2010; Jones, 2009; Navas-Castillo, Fiallo-Olivé, & Sánchez-Campos, 2011). A prominent example is the spread of Pepino mosaic virus (PepMV; genus *Potexvirus*, family *Alphaflexiviridae*), which was first described infecting *Solanum muricatum* plants in Peru in the 1970s (Jones, Koenig, & Lesemann, 1980) and emerged in tomato crops in Europe during the late 1990s (van der Vlugt et al., 2002). PepMV is causing worldwide outbreaks that are compromising tomato fruit quality and yields, resulting in important economic losses

*These authors contributed equally to this work.

  

(Gómez, Sempere, & Aranda, 2012; Hanssen & Thomma, 2010). There are no PepMV-resistant tomato cultivars that are commercially available yet, and despite the implementation of permanent surveillance and prevention programmes, PepMV represents a serious threat to tomato crops in many temperate regions of America, Africa, Asia and Europe (Gómez, Sempere, & Aranda, 2012). Five PepMV strains have been identified to date: (a) the Peruvian (LP) strain, originally found infecting *S. muricatum* and wild *Solanum* spp., (b) the European EU-tomato (EU) strain, (c) the American US1 strain, (d) the Chilean-2 (CH2) strain, and (e) the PES strain, recently described in wild tomato populations in Peru (Moreno-Pérez et al., 2014). Based on the genome sequence diversity of the PepMV populations sampled in cultivated tomato crops, the CH2 strain shares approximately 78% of their nucleotide identity with the rest of strains, while the nucleotide sequence of EU is notably similar to LP (90%) and closely related to US1 (86%) (Moreno-Pérez et al., 2014). After the first report of PepMV in Europe, PepMV populations were shown to be genetically very homogeneous in Spain, England and France, with isolates belonging to the EU strain (Cotillon, Girard, & Ducouret, 2002; Mumford & Metcalfe, 2001; Pagán et al., 2006), suggesting a single viral introduction from the Americas. Since 2005, this circumstance changed, due to the emergence of the US1 and CH2 strains (Alfaro-Fernández, Cebrián, Córdoba-Sellés, Herrera-Vásquez, & Jordá, 2008; Alfaro-Fernández, Sánchez-Navarro, del, & Cordoba, 2009; Gómez, Sempere, Elena, & Aranda, 2009; Hanssen et al., 2008; Hanssen & Thomma, 2010; Hasiów-Jaroszewska, Borodynko, Jackowiak, Figlerowicz, & Pospieszny, 2010; Ling, Wintermantel, & Bledsoe, 2008). Currently, it seems that PepMV-CH2 isolates have spread and become dominant, although the occurrence of isolates from other strains in mixed infections has also been reported (Gómez et al., 2009; Hanssen et al., 2008). This new context shows the need for further insights into the genetic diversity and structure of PepMV populations in those tomato-producing areas and appropriate management programmes.

PepMV is a potexvirus whose genome consists of a positive-sense, single-stranded RNA molecule, approximately 6.4 kb in length, containing five open reading frames (ORFs) flanked by two untranslated regions (UTRs), with a poly(A) tail at the 3′ end of the genomic RNA (Aguilar, Hernández-Gallardo, Cenis, Lacasa, & Aranda, 2002; Cotillon et al., 2002; Mumford & Metcalfe, 2001). ORF1 encodes the putative viral RNA-dependent RNA polymerase (RdRp), which contains three well-conserved domains found in the replicases from the other potexviruses: (a) the putative methyltransferase domain, (b) the NTPase/helicase domain containing NTP-binding motifs and (c) the RdRp domain. ORFs 2, 3, and 4 encode the triple gene block (TGB) proteins TGB1, TGB2 and TGB3, which are essential for virus movement, suppression of RNA silencing and building of the viral factory (Morozov & Solovyev, 2003; Verchot-Lubicz, 2005; Verchot-Lubicz, Ye, & Bamunusinghe, 2007). ORF5 encodes the coat protein (CP) that, in addition to its structural role, is required for cell-to-cell and long distance movement, and also plays a role in suppression of RNA silencing (Mathioudakis et al., 2012; Sempere, Gomez, Truniger, & Aranda, 2011). In terms of symptom expression, PepMV infection of tomato plants ranges from symptomless, mild and severe mosaic to necrosis, and in tomato fruits it ranges from marbling and blotchy ripening to malformation; symptom severity may be dependent on the viral strain, tomato cultivar and environment conditions (Sempere et al., 2015; Spence, Basham, Mumford, & Hayman, 2006).

PepMV is mechanically transmitted, and hence, cultural practices during fruit pruning and harvesting with contaminated tools, hands and clothing can easily lead to short-distance dispersal (Hasiów-Jaroszewska et al., 2010). Also, bumblebee pollinator species have been identified as potential vectors of PepMV (Lacasa et al., 2003; Shipp et al., 2008), and the high density of zoospores from the fungus *Olpidium virulentus* in the drainage water may increase PepMV transmission by irrigation (Schwarz, Beuch, Bandte, & Fakhro, 2010). Additionally, PepMV can be transmitted by vegetative propagation from infected tuber potato accessions (van der Vlugt, 2009). In terms of long-distance dispersal, seed transmission may constitute a major source of PepMV dispersion (Gómez, Sempere, & Aranda, 2012; Hasiów-Jaroszewska et al., 2010), and also marketing of tomato fruits may play a part in long-distance PepMV dispersal by the movement of infected fruits.

The first analysis of the natural genetic diversity of PepMV in Spain showed a rather genetically-uniform population, suggesting a single introduction (Pagán et al., 2006), although isolates of the PepMV-EU and US1 strains had already been detected in the Canary Islands in earlier epidemics (Alfaro-Fernández et al., 2008; Alfaro-Fernández et al., 2010), indicating separate PepMV introduction events associated with the islands' geographic isolation. Later, PepMV populations sampled in the Murcia region of south-eastern Spain between 2005 and 2008 were shown to comprise isolates of two co-circulating strains, CH2 and EU, and although the CH2 isolates predominated, the EU isolates persisted in mixed infections (Gómez et al., 2009). The presence of PepMV in Morocco has recently been demonstrated through the interception of infected tomato fruits; although no crop surveys have been conducted yet, the genetic characterisation of twelve isolates from fruits showed a high sequence identity with the CH2 strain (Souiri, Zemzami, Laatiris, Amzazi, & Ennaji, 2017).

In this study, we sequenced 85 isolates collected from eight major tomato-producing areas in Spain and Morocco during the 2017 year, allowing the description of the genetic diversity and structure of the PepMV populations in these areas. Both Spain and Morocco are among the largest tomato producers in the Mediterranean basin.

## 2 | MATERIALS AND METHODS

### 2.1 | Sample collection

A total of 231 apical leaf samples from tomato plants showing PepMV-like symptoms of infection were collected from different tomato crops grown in greenhouses (Table S1, Supporting information). Samples were collected from Spain (Murcia, Almería and Granada), Canary Islands (Tenerife and Gran Canaria) and Morocco (Agadir and Casablanca) in 2017 (Table S1 and Figure 1).

### 2.2 | PepMV detection

Plant samples were screened for PepMV infection by non-isotopic dot-blot hybridization analysis (Gómez et al., 2009). Molecular


**FIGURE 1** Map showing the geographic sources of PepMV isolates in Spain and Morocco

hybridization was carried out from total RNA extracts from tomato plant leaves, which were each blotted onto three positively-charged nylon membranes (Amersham Pharmacia Biotech), including samples of healthy plants and positive controls. Membrane were irradiated with UV light in a cross-linker, and then RNA:RNA hybridization detection was performed using three specific RNA probes that allowed for the simultaneous detection of the four different strains that have been reported in cultivated tomato. The EU_505 probe was complementary to nucleotides 1,388–1,711 (PepMV-EU) (Gómez et al., 2009), and is able to detect isolates from both EU and LP strains, and likewise isolates from CH2 and US1 strains were detected by the CH2_505 and US1_505 probes that were complementary to nucleotides 1,411–1891 (PepMV-CH2) and 1,408–1850 (PepMV-US1), respectively (Figure S1). These RNA probes were synthesised by transcription with SP6 RNA polymerase (New England Biolabs, Ipswitch, MA) from pGEM-T vectors (Promega corporation, Madison, WI) with the corresponding cDNA inserts. Molecular hybridization was carried out with 2 h of pre-hybridization and 68°C overnight incubation with the dig-labelled probe (Más & Pallás, 1995). Immunological detection using anti-dig antibody conjugated to phosphatase alkaline (Roche Diagnostics GmbH, Mannhein, Germany) was carried out after hybridization and stringency washes, and the chemiluminescent substrate CDP-Star (GE Healthcare UK Ltd, Buckinghamshire, England) and an Amersham Imager 600 (GE Healthcare Bio-Sciences AB, Uppsala, Sweden) was used to examine the membranes and detect PepMV.

## 2.3 | Cloning and sequencing analysis

Total RNA was extracted from a random collection of 85 PepMV-positive samples by using a NucleoSpin® RNA plant kit (Macherey-Nagel, GmbH & Co., Düren, Germany) and stored at −20°C until use. For each sample, one microgram of total RNA was used for amplification of the region encoding TGB1-3 and CP by RT-PCR, using Reverse Transcriptase (Roche) and Expand™ High Fidelity PCR System (Roche). The amplification was carried out following the manufacturer's instructions, using 50°C as the annealing temperature. RT-PCR products were cleaned-up and visualised by 0.7% agarose electrophoresis; all samples generated amplification products of approximately 2.2 kb, as expected. Each purified PCR fragment was cloned into a pGEM®-T Easy vector following the manufacturer's instructions. Stellar™ competent *Escherichia coli* cells (Takara Bio USA, Inc.) were transformed, and after plasmid verification based on the insert size, two random cDNA clones per sample were selected, except for two cases; TE17-1d and TE17-2d, for which three cDNA clones were selected. All of them were sequenced with the Sanger method in an external custom service (STAB vida, Caparica, Portugal).

## 2.4 | Phylogenetic and population structure analysis

The phylogenetic relationships between PepMV isolates were inferred from a collection of 13 PepMV sequences retrieved from GenBank (and referenced in this study according to its NCBI accession number), and the 172 sequences that were determined in this work. A total of 185 sequences were aligned by using the Multiple Sequence Comparison by Log-Expectation (MUSCLE) with a total of 1966 nucleotide positions in the final data set. The evolutionary history was inferred by using the Maximum Likelihood method based on the Tamura-Nei model run in MEGA7 (Kumar, Stecher, & Tamura, 2016). Initial tree(s) for the heuristic search were obtained automatically by applying the Neighbour Joining and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. Trees were drawn to scale and displayed using FigTree, with branch lengths measuring the number of substitutions per site. Estimates of genetic diversity for the populations of each geographical area and PepMV type were calculated using the DnaSP v6 software (Rozas et al., 2017). The analysis of potential recombination between PepMV isolates was carried out by using the methods implemented in RDP4 and a *P*-value of 0.05 (Martin, Murrell, Golden, Khoosal, & Muhire, 2015). Sub-populations among CH2 and EU types were examined using model-free discriminant analysis (DA) of principal components (DAPC) implemented in the R package adegenet (Jombart, Devillard, & Balloux, 2010). DAPC maximises variation between groups while minimising variation within groups. DAPC first transforms the data using principal component analysis (PCA) and then assigns individuals to clusters by DA. The data transformation ensures that the input of variables to DA are uncorrelated and that their number is less than the number of individuals analysed, so as to overcome the drawbacks of direct application of DA (Wang et al., 2017). Genetic distances among variants of the CH2 and EU types were visualised by a minimum-spanning network (MSN), using the function *poppr.msn* implemented in the R poppr package (Kamvar, Brooks, & Grunwald, 2015).

## 2.5 | Data availability

All sequencing information and data that support the findings of this study have been deposited in GenBank under accession numbers MH550173–MH550344.

## 3 | RESULTS

### 3.1 | PepMV strain prevalence

We examined 231 samples from symptomatic tomato plants from 30 tomato varieties obtained in a survey carried out in eight localities of Spain and Morocco (Figure 1) during the year 2017. A total of 198 samples were PepMV-positive, and 33 samples were either non-infected or had an undetectable PepMV load. PepMV was detected in all surveyed locations (Table 1; additional information is summarised in Table S1). In particular, results using specific RNA probes to distinguish PepMV strains showed that most of the PepMV isolates belonged to the PepMV-CH2 strain (86% of the infected samples), being the only PepMV strain found in Águilas and Morocco. The PepMV-EU strain was detected in 34% of the infected samples, while the PepMV-US1 strain was only found in three infected samples from Tenerife. Notably, isolates of the PepMV-EU strain were mainly found in mixed infections with isolates of the CH2 strain, although PepMV-EU single infections were detected in Gran Canaria South (100% of the Gran Canaria South samples), Tenerife (4.3% of the Tenerife samples) and Almería (4.8% of the Almería samples) (Table 1).

### 3.2 | PepMV genetic variability and population structure

We cloned the genomic fragment that includes the triple gene block (TGB1, TGB2 and TGB3) and the CP gene of 85 random PepMV-positive samples (Table S1) based on location, and for most of them two cDNA clones were sequenced per sample. After sequence alignment, a matrix of genetic distances between PepMV isolates was computed to build a maximum likelihood tree for each ORF independently (data not shown) and for all concatenated coding sequences (Figure 2). PepMV isolates clearly clustered into three genetic groups that corresponded with three PepMV strains (CH2, EU and US1), in agreement with the above hybridization results. Within each cluster, sequences showed a high genetic similarity, although some sequences appeared to be geographically grouped within clusters (Figure 2). We also searched for potential recombinant sequences (data not shown), and no recombinants were identified. We then sought to test any subdivision among the PepMV populations, considering a population as the isolates belonging to each strain. To this end, we first used Discriminant Analyses of Principal Components (DAPC), and found that the PepMV-CH2 strain population was discriminated into three sub-populations (Figure 3a) which were associated with the geographical distance, Peninsular Spain, Gran Canaria and Morocco. This DAPC biplot showed a gradient in the first discriminant component that separated the Gran Canarian CH2 isolates, while the second discriminant component separated the Moroccan ones from the rest of the isolates. It should be noted that despite the geographical proximity between the Tenerife and Gran Canaria islands, the isolates from Tenerife were closer to those from Peninsular Spain than to the Gran Canaria ones. Although the number of PepMV-EU isolates examined was small, which could result in biased results, we also evaluated the PepMV-EU strain population; DAPC discriminated four subpopulations (Figure 4a). The DAPC biplot showed that the first discriminant component separated EU isolates from Tenerife and Granada, leaving those from Mazarrón and Almería together, while the second discriminant component separated isolates from Gran Canaria. The geographical subdivisions could be suggesting different PepMV dispersal pathways (short and long-distance) between tomato producing-areas.

**TABLE 1** Plant tomato samples analysed in this study, grouped by country and location

| Country | Location | Total samples/PepMV-infected samples | PepMV strain[a] | | | |
|---|---|---|---|---|---|---|
| | | | EU | CH2 | US1 | Mixed |
| Spain | Mazarrón | 48/35 | 15 | 32 | 0 | 12 |
| | Águilas | 28/24 | 0 | 24 | 0 | 0 |
| | Almería | 25/21 | 2 | 20 | 0 | 1 |
| | Granada | 23/22 | 3 | 22 | 0 | 3 |
| | Gran Canaria (north) | 18/18 | 15 | 14 | 0 | 11 |
| | Gran Canaria (east) | 12/8 | 0 | 8 | 0 | 0 |
| | Gran Canaria (south) | 18/17 | 17 | 0 | 0 | 0 |
| | Gran Canaria (west) | 12/7 | 0 | 8 | 0 | 0 |
| | Tenerife | 23/23 | 16 | 20 | 3 | 15 |
| | Casablanca | 8/8 | 0 | 8 | 0 | 0 |
| Morocco | | | | | | |
| | Agadir | 16/15 | 0 | 15 | 0 | 0 |
| | | **231/198** | **68** | **171** | **3** | **42** |

[a] Total RNA from each sample was analysed by dot-blot hybridization using specific probes for the EU, CH2 and US1 strain detection.


**FIGURE 2** Phylogenetic analysis of PepMV isolates. Maximum likelihood tree of 172 sequences from 85 PepMV isolates collected in tomato-producing areas from Spain and Morocco in 2017, including 13 reference sequences from GenBank and referenced according to its NCBI accession number. All sequences were aligned with MUSCLE and the evolutionary history was inferred by using the maximum likelihood method based on the Tamura-Nei model performed in MEGA7 (Kumar et al., 2016). The tree branch lengths measure the number of substitutions per site, and the scale bar represents genetic distance. The PepMV-CH2 clade is shown in red colour, PepMV-EU in green colour and PepMV-US1 in dark yellow colour. The nomenclature of each isolate indicates the location (MZ, Mazarrón; AG, Águilas, AL, Almería, GR, Granada; GC, Gran Canaria; TE, Tenerife; MA, Morocco), year (17), sample-plot (n°1–4), and sample-code (a-h) (see Table S1 for details)

### 3.3 | Linking genetic diversity and PepMV spatial distribution

In order to obtain further insights into the role of the geographical PepMV population differentiation, we estimated the genetic diversity ($\pi$), the interpopulation genetic differentiation between viral populations (Fst) and the population demography or tests of neutrality (based on Tajima's D, Fu and Li's D and Fu and Li's F values) using the sequencing data from PepMV populations. The genetic diversity for the entire population was very low ($\pi$ = 0.006) with a moderate gene flow between populations (Fst = 0.22), and no significant values in neutrality tests (Tajima's D, Fu and Li's D, and Fu and Li's F). However, the differential strain prevalence among locations could be masking subtle effects; the number of segregating sites for locations with isolates belonging to only one strain was very low (Águilas and Morocco, with 87 and 100 segregating sites, respectively) compared to the rest of the populations (which ranged between 459 and 566 segregating sites). We therefore re-analysed the above genetic parameters according to strain (CH2 and EU) (Table 2).

Focusing on the PepMV-CH2 strain, the analysis of the diversity of the four genomic regions displayed low nucleotide diversity values, ranging from 0.011 to 0.014, and tests significantly deviated from neutrality ($p < 0.001$) (Table 2). Thus, the diversity in synonymous (dS) and non-synonymous (dN) positions was calculated to infer the strength of the selection pressure (Table 2); no site was detected under positive selection for any gene, whereas several sites were detected under negative selection for genes encoding TGB1 and CP (data not shown). These data indicated strong purifying selection, which was clearly supported by low dN/dS values (Table 2). We then performed a MSN analysis to visualise genetic relatedness among PepMV-CH2 isolates (Figure 3b). The genetic differentiation of the Moroccan isolates was clear. Another small group was composed of Gran Canaria and Tenerife isolates that were in turn differentiated among themselves. The Spanish Peninsular isolates appeared dispersed between the two main groups (Figure 3b), suggesting gene flow between crops among Spanish Peninsular locations. Thereafter, we estimated Fst values to determine the extent of the gene flow or migration among PepMV-CH2 populations, and found that these values ranged from 0.05 to 0.35, which suggested a moderate to frequent gene flow among populations. Relatively moderate gene flow was inferred between Morocco and Tenerife (Fst = 0.35) and Morocco and Granada (Fst = 0.26), followed by Tenerife and Águilas (Fst = 0.23), and frequent gene flow was identified between Almería and Granada (Fst = 0.13), Almería and Águilas (Fst = 0.15), and the higher rate of gene flow was between Almería and Mazarrón

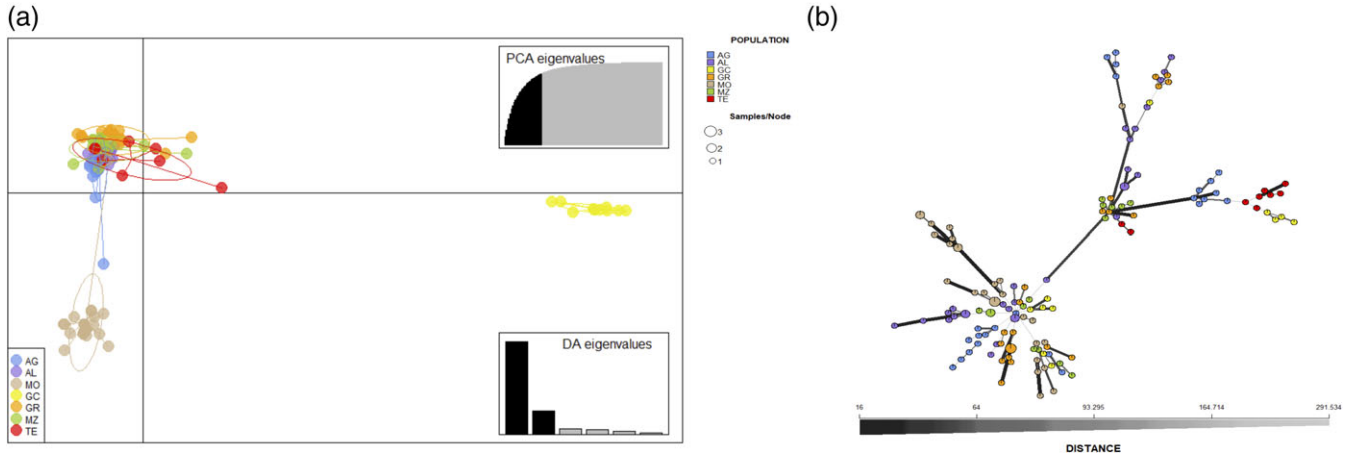


**FIGURE 3** (a) Discriminant principal components analysis (DAPC) clusters of PepMV-CH2 isolates from Spain and Morocco. AG, Águilas; AL, Almería; MA, Morocco; GC, Gran Canaria; GR, Granada; MZ, Mazarrón; TE, Tenerife. Eigenvalues retained for principal component and discriminant functions in the analysis are displayed in insets. (b) Minimum spanning network of the PepMV-CH2 genotype variants from Spain and Morocco. MZ, Mazarrón; AG, Águilas; AL, Almería; GR, Granada; GC, Gran Canaria; TE, Tenerife; MA, Morocco. Each node represents a different variant, and node sizes and colours correspond to the number of individuals and geographic population, respectively. Edge width and shading represent relatedness in accordance to the genetic distance scale, which is displayed in inset

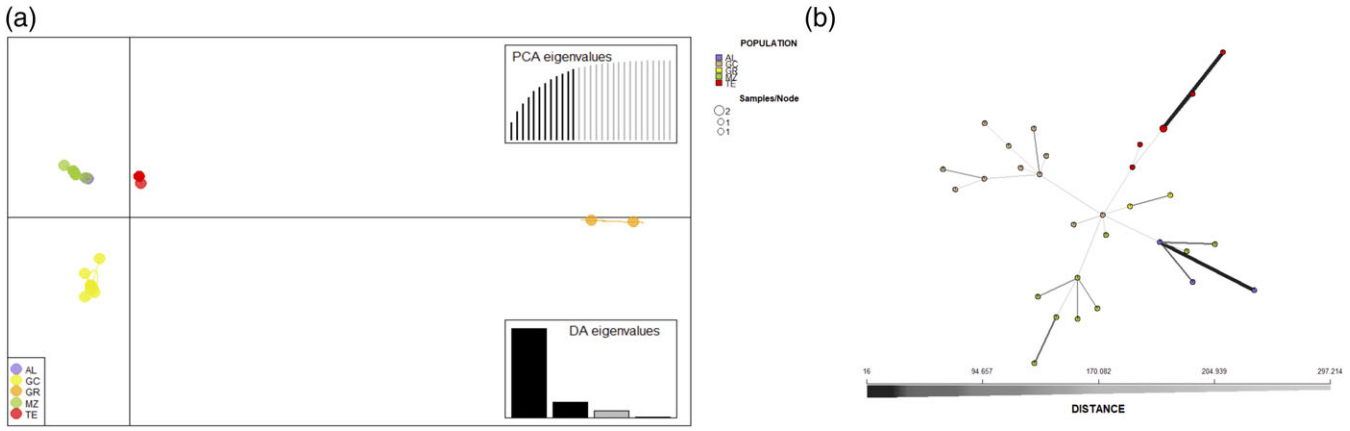


**FIGURE 4** (a) Discriminant principal components analysis (DAPC) cluster of PepMV-EU isolates from Spanish populations. AL, Almería; GC, Gran Canaria; GR, Granada; MZ, Mazarrón; TE, Tenerife. Eigenvalues retained for principal component and discriminant functions in the analysis are displayed in inset. (b) Minimum spanning network of the PepMV-EU genotype variants from Spanish populations. Each node represents a different variant, and node sizes and colours correspond to the number of individuals and geographic population, respectively. Edge width and shading represent relatedness in accordance to the genetic distance scale, which is displayed in insets

(Fst = 0.05). These results show that despite the low genetic variability among PepMV-CH2 populations, geographic differences could be identified, probably as a result of different gene flow rates.

As for the PepMV-EU populations, these also showed low nucleotide diversity values, ranging from 0.013 to 0.016 (average $\pi$ = 0.015) for the four coding regions, with synonymous and non-synonymous diversity values comparable to those of PepMV-CH2 populations. Neutrality tests produced negative values, and only deviated from neutrality for the TGB1 gene (Table 2), with 12 sites under negative selection. The MSN analysis of the PepMV-EU population showed geographical differentiation between the Tenerife, Gran Canaria and Mazarrón populations (Figure 4b). This result was in agreement with the high genetic variation in allele frequencies identified across populations (Fst = 0.56), which indicated a low gene flow frequency.

## 4 | DISCUSSION

This study provides insights on the occurrence and molecular characterisation of PepMV in two major tomato-producing areas in Spain (south-eastern Spain and Canary Islands) and Morocco in 2017. Most of the samples (86%) from plants exhibiting symptoms of PepMV-like infection resulted positive for PepMV by dot-blot hybridization. Also, taking into account that PepMV may comprise viral variants that range from aggressive to asymptomatic, and symptom severity can vary considerably due to environmental factors, a relatively negligible number of tomato symptomatic samples (14%) were either non-infected by PepMV or PepMV was undetected in them. This may be due to either a misdiagnosis of the symptomatology due to any plant disorders associated with chemicals use in farming practices or as a consequence of the infection status in the leaves at the time of sampling, although we cannot rule out the presence of other viruses. Our

**TABLE 2** Genetic diversity parameters (± SE) and neutrality tests estimated for the triple gene block proteins TGB1, TGB2 and TGB3, and the coat protein (CP) of the PepMV-CH2 and -EU strains

| PepMV strain | Genomic region | n | S | π | dS | dN | dN/dS | Tajima's D | D*Fu & Li | F*Fu & Li | Ne | Po |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **CH2** | TGBp1 | 701 | 0.22 | 0.013 | 0.038 ± 0.006 | 0.005 ± 0.001 | 0.13 | −2.23180* | −2.19199 | −2.64824* | 23 | 0 |
| | TGBp2 | 371 | 0.20 | 0.011 | 0.028 ± 0.005 | 0.003 ± 0001 | 0.11 | −2.25793* | −2.82477* | −3.09421* | 5 | 0 |
| | TGBp3 | 248 | 0.20 | 0.012 | 0.021 ± 0.002 | 0.009 ± 0.002 | 0.43 | −2.06480* | −3.40542* | −3.40985* | 3 | 0 |
| | CP | 711 | 0.26 | 0.014 | 0.040 ± 0004 | 0.005 ± 0001 | 0.12 | −2.38746* | −3.17532* | −3.36973* | 34 | 0 |
| **EU** | TGBp1 | 704 | 0.10 | 0.013 | 0.023 ± 0.005 | 0.008 ± 0.002 | 0.35 | −1.88912* | −2.55541* | −2.75957* | 4 | 0 |
| | TGBp2 | 371 | 0.10 | 0.014 | 0.024 ± 0.006 | 0.007 ± 0.002 | 0.29 | −1.61229 | −0.94024 | −1.37056 | 0 | 0 |
| | TGBp3 | 254 | 0.10 | 0.016 | 0.022 ± 0.008 | 0.013 ± 0.004 | 0.59 | −1.31322 | −0.66325 | −1.02906 | 0 | 0 |
| | CP | 712 | 0.11 | 0.016 | 0.043 ± 0.006 | 0.006 ± 0.002 | 0.14 | −1.67797 | −1.91077 | −2.16896 | 8 | 0 |

π: nucleotide diversity (mean nucleotide differences per site between sequence pairs); dS: frequency of synonymous substitution per site; dN: frequency of nonsynonymous substitution per site; D Tajima: D*Fu & Li; F*Fu & Li statistics tests of neutrality employed in population demography analyses; *n*: number of sites; Ne: number of negatively selected codons; Po: number of positively selected codons; S: number of segregating (polymorphic) sites.
*$p < 0.05$.

study focused on the PepMV-positive samples and showed that Occidental Mediterranean PepMV populations are predominantly composed of PepMV isolates belonging to the CH2 strain, although isolates of the EU strain were also detected in Spain in significant proportions, particularly in Tenerife, Gran Canaria and Mazarrón. A few isolates of the US1 strain were also detected in Tenerife, and it should be noted that a few more EU isolates were detected in single infections in three plots from the south of Gran Canaria, where most of EU detections were found in mixed infections, as reported before (Gómez et al., 2009). In turn, and despite the identification of recombinant RdRp sequences in Belgian greenhouse tomatoes (Hanssen et al., 2008), recombinant strains have not been detected in Spanish populations, even after the lengthy occurrence of mixed infections in Spanish tomato crops (Gómez et al., 2009; Gómez, Sempere, & Aranda, 2012), suggesting that potential recombinants of different PepMV isolates may not have been favoured in the populations, although further research would be required to clarify this issue. In contrast, Moroccan and Águilas (south-eastern Spain) populations were only composed of CH2 strain isolates, which is particularly remarkable for the Águilas population, given its proximity to Mazarrón. The prevalence of the CH2 strain agrees with previous studies on the PepMV population structure in Spain (Gómez et al., 2009; Gómez, Sempere, & Aranda, 2012), Belgium (Hanssen et al., 2008; Hanssen & Thomma, 2010), Poland (Hasiów-Jaroszewska et al., 2010) and North-America (Ling, Li, & Bledsoe, 2013).

The phylogenetic and genetic differentiation analyses carried out showed that PepMV genetic variability was low within the EU and CH2 strains, and comparable to those previously found in Spain (Gómez et al., 2009; Pagán et al., 2006). Indeed, our neutrality tests showed negative values and deviated significantly from neutrality (Table 2) with several sites in the TGB1 and CP genes under negative selection, which was consistent with the above genetic diversity results as well as the purifying selection identified in previous studies (Gómez et al., 2009). The negative values obtained for the PepMV population indicated an excess of single-nucleotide variants with low frequency in the population (Blanco-Meneses, Carbone, & Ristaino, 2018), and this could be suggesting that PepMV populations, in particular those of the CH2 strain, were in demographic expansion. Our results also showed frequent gene flow that contributes to genetically homogenise sub-populations that are geographically close, revealing the importance of gene flow in shaping the genetic structure and evolutionary dynamics of PepMV. Nevertheless, the PepMV populations appeared to be geographically differentiated; PepMV-CH2 populations differed between geographically-distant populations (Figure 3a), and this could also be the case for the PepMV-EU populations, although the relative low number of isolates makes this a rather preliminary conclusion. Thus, the genetic diversity of the PepMV population could be explained by purifying selection and moderate gene flow within subpopulations. Furthermore, agro-ecological factors could contribute with the increasing of the effective population size (Gómez, Sempere, Aranda, & Elena, 2012) favouring selection to operate over population differentiation.

The existence of PepMV-CH2 in Morocco was shown for the first time in 2016 by the interception of infected tomato fruits (Souiri et al., 2017). Here we showed the first molecular identification of Moroccan PepMV populations through a crop survey. Our minimum spanning network analysis showed that the Moroccan isolates were related to other Spanish isolates (Figure 3b), suggesting their introduction from Peninsular Spain. However, this pattern may also indicate that independent introductions of PepMV occurred in Spain and Morocco and have gone unnoticed in Morocco until now; the global commerce of tomato fruits and PepMV seed transmission may be contributing to long-distance PepMV genotype dispersion. Broader phylogeographic analyses would be useful to test this hypothesis.

In conclusion, the PepMV population structure in Spain and Morocco showed that the CH2 strain prevails, although the EU strain was also found predominantly under mixed infections, and the US1 genotype was restricted to a few cases in Tenerife. In general, our results support the finding of low genetic diversity, but geographical differentiation was counteracted by moderate gene flow. This knowledge could provide insights into fundamental eco-evolutionary processes and valuable information for the design and implementation of stable and effective control strategies, such as cross-protection by attenuated PepMV isolates (Agüero et al., 2018; Chewachong et al., 2015; Hanssen et al., 2010), which appears to be a good solution in the short term for managing this viral disease.

## ACKNOWLEDGEMENTS

We thank the assistance of practitioners and technicians from each tomato-producing area, and also to the anonymous reviewers for the detailed and helpful comments to the manuscript. C.A. was supported by funding of the Ministry of Economy, Industry and Competitiveness (MINECO, Spain) within a PhD programme grant (FPU16/02569). C.G.A. and J.A. were recipients of grants PTQ-15-07646 and PTQ-13-05882, respectively, from the Torres-Quevedo program (Ministry of Economy, Industry and Competitiveness; Spain). F.M. was supported by funding from the 'Ministère de l'Enseignement Supérieur, de la Recherche Scientifique de la Formation des Cadres' (MERSFC, Morocco) within the framework of ARIMNet2 Project. P.G. and M.A.A. also acknowledge funding from MINECO (Spain) from research grants AGL2014-59556-R and PCIN-2017-055 (ERANET-ARIMNet2, ref. 302), respectively. We also thank Mario Fon (mariogfon@gmail.com) for editorial assistance.

## CONFLICTS OF INTEREST

The authors declare that there is no conflict of interest.

## ORCID

*Pedro Gómez* https://orcid.org/0000-0003-2830-4105

## REFERENCES

Agüero, J., Gómez-Aix, C., Sempere, R. N., García-Villalba, J., García-Núñez, J., Hernando, Y., & Aranda, M. A. (2018). Stable and broad spectrum cross-protection against Pepino mosaic virus attained by mixed infection. *Frontiers in Plant Science*, *9*, 1810. https://doi.org/10.3389/fpls.2018.01810

Aguilar, J. M., Hernández-Gallardo, M. D., Cenis, J. L., Lacasa, A., & Aranda, M. A. (2002). Complete sequence of the *Pepino mosaic virus* RNA genome. *Archives of Virology*, *147*, 2009–2015.

Alfaro-Fernández, A., Cebrián, M. C., Córdoba-Sellés, C., Herrera-Vásquez, J. A., & Jordá, C. (2008). First report of the US1 strain of *Pepino mosaic virus* in tomato in the Canary Islands, Spain. *Plant Disease*, *92*, 1590–1591.

Alfaro-Fernández, A., Cebrián, M. C., Herrera-Vásquez, J. A., Córdoba-Sellés, M. C., Sánchez-Navarro, J. A., & Jordá, C. (2010). Molecular variability of Spanish and Hungarian isolates of *Tomato torrado virus*. *Plant Pathology*, *59*, 785–793.

Alfaro-Fernández, A., Sánchez-Navarro, J. Á., del Carmen Cebrián, M., del Carmen Córdoba-Sellés, M., Pallás V., & Jordá, C. (2009). Simultaneous detection and identification of *Pepino mosaic virus* (PepMV) isolates by multiplex one-step RT-PCR. *European Journal of Plant Pathology*, *125*, 143–158.

Anderson, P. K., Cunningham, A. A., Patel, N. G., Morales, F. J., Epstein, P. R., & Daszak, P. (2004). Emerging infectious diseases of plants: Pathogen pollution, climate change and agrotechnology drivers. *Trends in Ecology & Evolution*, *19*, 535–544.

Blanco-Meneses, M., Carbone, I., & Ristaino, J. B. (2018). Population structure and migration of the tobacco blue mold pathogen, *Peronospora tabacina*, into North America and Europe. *Molecular Ecology*, *27*, 737–751.

Chewachong, G. M., Miller, S. A., Blakeslee, J. J., Francis, D. M., Morris, T. J., & Qu, F. (2015). Generation of an attenuated, cross-protective *Pepino mosaic virus* variant through alignment-guided mutagenesis of the viral capsid protein. *Phytopathology*, *105*, 126–134.

Cotillon, A. C., Girard, M., & Ducouret, S. (2002). Complete nucleotide sequence of the genomic RNA of a French isolate of *Pepino mosaic virus* (PepMV). *Archives of Virology*, *147*, 2231–2238.

Gómez, P., Sempere, R., Aranda, M., & Elena, S. (2012). Phylodynamics of *Pepino mosaic virus* in Spain. *European Journal of Plant Pathology*, *134*, 445–449. https://doi.org/10.1007/s10658-012-0019-0

Gómez, P., Sempere, R. N., & Aranda, M. A. (2012). Chapter 14 – *Pepino mosaic virus* and *Tomato Torrado virus*: Two emerging viruses affecting tomato crops in the Mediterranean Basin. In G. Loebenstein & H. Lecoq (Eds.), *Advances in virus research 84: Viruses and virus diseases of vegetables in the Mediterranean Basin*. (pp. 505–532). Amsterdam: Academic Press.

Gómez, P., Sempere, R. N., Elena, S. F., & Aranda, M. A. (2009). Mixed infections of *Pepino mosaic virus* strains modulate the evolutionary dynamics of this emergent virus. *Journal of Virology*, *83*, 12378–12387.

Hanssen, I. M., Gutierrez-Aguirre, I., Paeleman, A., Goen, K., Wittemans, L., Lievens, B., … Thomma, B. P. H. J. (2010). Cross-protection or enhanced symptom display in greenhouse tomato co-infected with different *Pepino mosaic virus* isolates. *Plant Pathology*, *59*, 13–21.

Hanssen, I. M., Lapidot, M., & Thomma, B. (2010). Emerging viral diseases of tomato crops. *Molecular Plant-Microbe Interactions*, *23*, 539–548.

Hanssen, I. M., Paeleman, A., Wittemans, L., Goen, K., Lievens, B., Bragard, C., … Thomma, P. H. J. (2008). Genetic characterization of *Pepino mosaic virus* isolates from Belgian greenhouse tomatoes reveals genetic recombination. *European Journal of Plant Pathology*, *121*, 131–146.

Hanssen, I. M., & Thomma, B. (2010). *Pepino mosaic virus*: A successful pathogen that rapidly evolved from emerging to endemic in tomato crops. *Molecular Plant Pathology*, *11*, 179–189.

Hasiów-Jaroszewska, B., Borodynko, N., Jackowiak, P., Figlerowicz, M., & Pospieszny, H. (2010). *Pepino mosaic virus* – A pathogen of tomato crops in Poland: Biology, evolution and diagnostics. *Journal of Plant Protection Research*, *50*, 470–476.

Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genetics*, *11*, 94.

Jones, R. A. C. (2009). Plant virus emergence and evolution: Origins, new encounter scenarios, factors driving emergence, effects of changing world conditions, and prospects for control. *Virus Research*, *141*, 113–130.

Jones, R. A. C., Koenig, R., & Lesemann, D. E. (1980). *Pepino mosaic virus*, a new potexvirus from pepino (*Solanum muricatum*). *Annals of Applied Biology*, *94*, 61–68.

Kamvar, Z. N., Brooks, J. C., & Grunwald, N. J. (2015). Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Frontiers in Genetics*, *6*, 208.

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular and Biological Evolution*, *33*, 1870–1874.

Lacasa, A., Guerrero, M. M., Hita, I., Martínez, M. A., Jordá, C., Bielza, P., … Cano, A. (2003). Implication of bumble bees (*Bombus* spp.) on *Pepino mosaic virus* (PepMV) spread on tomato crops. *Plagas*, *29*, 393–403.

Ling, K.-S., Li, R., & Bledsoe, M. (2013). *Pepino mosaic virus* genotype shift in North America and development of a loop-mediated isothermal amplification for rapid genotype identification. *Virology Journal*, *10*, 117.

Ling, K.-S., Wintermantel, W. M., & Bledsoe, M. (2008). Genetic composition of *Pepino mosaic virus* population in North American greenhouse tomatoes. *Plant Disease*, *92*, 1683–1688.

Martin, D. P., Murrell, B., Golden, M., Khoosal, A., & Muhire, B. (2015). RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evolution*, *1*, vev003.

Más, P., & Pallás, V. (1995). Non-isotopic tissue-printing hybridization: A new technique to study long-distance plant virus movement. *Journal of Virological Methods*, *52*, 317–326.

Mathioudakis, M. M., Veiga, R., Ghita, M., Tsikou, D., Medina, V., Canto, C., … Livieratos, I. C. (2012). *Pepino mosaic virus* capsid protein interacts with a tomato heat shock protein cognate 70. *Virus Research*, *163*, 28–39.

Moreno-Pérez, M. G., Pagán, I., Aragón-Caballero, L., Cáceres, F., Fraile, A., & García-Arenal, F. (2014). Ecological and genetic determinants of *Pepino mosaic virus* emergence. *Journal of Virology*, *88*, 3359–3368.

Morozov, S. Y., & Solovyev, A. G. (2003). Triple gene block: Modular design of a multifunctional machine for plant virus movement. *Journal of General Virology*, *84*, 1351–1366.

Mumford, R. A., & Metcalfe, E. J. (2001). The partial sequencing of the genomic RNA of a UK isolate of *Pepino mosaic virus* and the comparison of the coat protein sequence with other isolates from Europe and Peru. *Archives of Virology*, *146*, 2455–2460.

Navas-Castillo, J., Fiallo-Olivé, E., & Sánchez-Campos, S. (2011). Emerging virus diseases transmitted by whiteflies. *Annual Review of Phytopathology*, *49*, 219–248.

Oerke, E. C. (2005). Crop losses to pests. *The Journal of Agricultural Science*, *144*, 31–43.

Pagán, I., Córdoba-Sellés, M., Martínez-Priego, L., Fraile, A., Malpica, J. M., Jordá, C., & García-Arenal, F. (2006). Genetic structure of the population of *Pepino mosaic virus* infecting tomato crops in Spain. *Phytopathology*, *96*, 274–279.

Rozas, J., Ferrer-Mata, A., Sanchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., & Sánchez-Gracia, A. (2017). DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular and Biological Evolution*, *34*, 3299–3302.

Schwarz, D., Beuch, U., Bandte, M., & Fakhro, A. (2010). Spread and interaction of *Pepino mosaic virus* (PepMV) and *Pythium aphanidermatum* in a closed nutrient solution recirculation system: Effects on tomato growth and yield. *Plant Pathology*, *59*, 443–452.

Sempere, R., Gomez, P., Truniger, V., & Aranda, M. (2011). Development of expression vectors based on *Pepino mosaic virus*. *Plant Methods*, *7*, 6.

Sempere, R. N., Gómez-Aix, C., Ruíz-Ramón, F., Gómez, P., Hasiów-Jaroszewska, B., Sánchez-Pina, M. A., & Aranda, M. A. (2015). *Pepino mosaic virus* RNA-dependent RNA polymerase POL domain is a hypersensitive response-like elicitor shared by necrotic and mild isolates. *Phytopathology*, *106*, 395–406.

Shipp, J. L., Buitenhuis, R., Stobbs, L., Wang, K., Kim, W. S., & Ferguson, G. (2008). Vectoring of *Pepino mosaic virus* by bumble-bees in tomato greenhouses. *Annals of Applied Biology*, *153*, 149–155.

Souiri, A., Zemzami, M., Laatiris, H., Amzazi, S., & Ennaji, M. M. (2017). First report of *Pepino mosaic virus* in tomato in Morocco. *New Disease Reports*, *35*, 10.

Spence, N. J., Basham, J., Mumford, R. A., & Hayman, G. (2006). Effect of *Pepino mosaic virus* on the yield and quality of glasshouse-grown tomatoes in the UK. *Plant Pathology*, *55*, 595–606.

van der Vlugt, R. A. A. (2009). *Pepino mosaic virus*. *Hellenic Plant Protection Journal*, *2*, 47–56.

van der Vlugt, R. A. A., Cuperus, C., Vink, J., Stijger, I. C. M. M., Lesemann, D.-E., Verhoeven, J. T. J., & Roenhorst, J. W. (2002). Identification and characterization of *Pepino mosaic potex virus* in tomato. *Bulletin OEPP*, *32*, 503–508.

Verchot-Lubicz, J. (2005). A new cell-to-cell transport model for *Potexviruses*. *Molecular Plant-Microbe Interactions*, *18*, 283–290.

Verchot-Lubicz, J., Ye, C. M., & Bamunusinghe, D. (2007). Molecular biology of potexviruses: Recent advances. *Journal of General Virology*, *88*, 1643–1655.

Vurro, M., Bonciani, B., & Vannacci, G. (2010). Emerging infectious diseases of crop plants in developing countries: Impact on agriculture and socio-economic consequences. *Food Security*, *2*, 113–132.

Wang, J., Fernandez-Pavia, S. P., Larsen, M. M., Garay-Serrano, E., Gregorio-Cipriano, R., Rodríguez-Alvarado, G., … Goss, E. M. (2017). High levels of diversity and population structure in the potato late blight pathogen at the Mexico Centre of origin. *Molecular Ecology*, *26*, 1091–1107.

Woolhouse, M. E., Haydon, D. T., & Antia, R. (2005). Emerging pathogens: The epidemiology and evolution of species jumps. *Trends in Ecological Evolution*, *20*, 238–244.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**How to cite this article:** Gómez-Aix C, Alcaide C, Agüero J, et al. Genetic diversity and population structure of *Pepino mosaic virus* in tomato crops of Spain and Morocco. *Ann Appl Biol*. 2019;174:284–292. https://doi.org/10.1111/aab.12491